マルコフ確率場のパラメータ推定アルゴリズムおよび高次マルコフ確率場に対する発展的な平均場近似法の開発

著者	高橋 茶子
学位授与機関	Tohoku University
学位授与番号	11301甲第19360号
URL	http://hdl.handle.net/10097/00130250

学位論文題目　　マルコフ確率場のパラメータ推定アルゴリズムおよび高次マルコフ確率場に対する発展的な平均場近似法の開発

氏名　　高橋　茶子

研究科、専攻　　東北大学大学院情報科学研究科（博士課程）応用情報科学専攻

論 文 内 容 の 要 約

第１章　序論

数理モデルを作ることは数理モデル化と呼ばれ、例えば、現象や構造の理解、時間や温度などの環境の変化による対象の状態の変化の解明および説明、観測を得るために高いコストが必要となる（もしくは現実的に不可能な）実験のシミュレーション、予測や意思決定、異常検知やシステム制御などの多岐にわたる分野で役立っている。これまでの数理モデル化の方法の発展は、人工知能技術の発展と密接な関わりを持っている。特に 2000 年以降はデータから重要な情報を抽出することで数理モデルを構築する機械学習の方法の発展が目覚ましい。データが何らかの生成規則から発生していると仮定し、データを用いてその生成規則を確率モデルで表現する機械学習の枠組みを統計的機械学習と呼ぶ。統計的機械学習による数理モデル化の汎用的な方法を確立することで、データを扱うあらゆる分野への応用の道が開けることから、統計的機械学習への関心は年々高まっている。

計算処理や理論的解釈の面で扱いやすいモデル表現を用いることは、モデル構築や数理モデルを利用した問題解決を容易にするために重要である。本論文では、そのような面で扱いやすい性質を持つマルコフ確率場を用いた統計的機械学習に着目する。データを用いた数理モデルの構築処理と構築されたモデルを用いた推論処理は明示的に分けて考えることができる。データの生成規則をあるモデルの形で記述するための一連の計算は学習と呼ばれる逆問題である。一方、学習により構築されたモデルに対してある入力を与えたとき、どのような出力が得られるのかを推定する計算は推論と呼ばれる順問題である。学習したモデルを用いて予測やシミュレーション、意思決定などを行うための計算処理は推論に該当する。本論文では、マルコフ確率場の学習と推論について、二つの個別の問題を検討する。

第２章　確率的グラフィカルモデルとマルコフ確率場

第２章では、第３章から第５章までの内容の基盤となる、確率的グラフィカルモデルおよびマルコフ確率場の基礎的な事項を説明する。確率的グラフィカルモデルとは、グラフを用いて確率変数間の関係を表すことができる確率モデルの総称である。グラフはノードおよびエッジで構成され、各ノードに確率変数が割り当てられる。ノード間のエッジは変数間の相互作用を表す。エッジに向きのない確率的グラフィカルモデルは無向グラフィカルモデルと呼ばれる。本章では、有向グラフィカルモデルを含む確率的グラフィカルモデル全般がさまざまな分野で用いられるようになってきた経緯について説明する。無向グラフィカルモデルのうち、条件付き独立性についての重要な性質であるマルコフ性を満たすモデルがマルコフ確率場である[Lauritzen, 1996]。

マルコフ性およびマルコフ性に密接に関連する性質である因数分解性について説明する。最後に、本論文の第３章以降で扱うペアワイズマルコフ確率場および主に第５章で扱う高次マルコフ確率場を定義する。ペアワイズマルコフ確率場とは二つの変数の間の相互作用をパラメータとして含むマルコフ確率場であり、高次マルコフ確率場とは n 個の変数の間の相互作用をパラメータとして含むマルコフ確率場である。

第３章　マルコフ確率場の相互作用推定問題の統計力学的解析

第３章では、マルコフ確率場の学習に焦点を当てる。インターネットの普及や計算機の処理能力の向上、ストレージの増大、計測機器の性能向上などにより、取得できるデータの量は飛躍的に増加した。これにより大量のデータから対象の特徴を抽出し数理モデル化を行う技術が多くの成功をおさめている。しかしながら、データが手に入りやすい状況に恵まれた領域ばかりが存在するわけではなく、技術的問題、時間的問題、経済的問題などのさまざまな要因により、データの取得が困難な問題が多く残っている。本章では、利用できるデータの数が推定すべきモデルパラメータの数よりも少ない、劣決定系における学習を考える。マルコフ確率場のエネルギー関数（ハミルトニアン）は、入力としてスピン配位が与えられた場合にエネルギー値を返す関数であるとみなすことができるため、データである入力と出力のペアからのマルコフ確率場のハミルトニアンの推定は、典型的な逆問題であると考えることができる。しかしながら、劣決定系と呼ばれるような状況で単純にパラメータの推定を行うことを考えると、パラメータの推定解は一意に定まらないため、何らかの工夫が必要である。本章では、信号処理の分野で発展している、圧縮センシング[Donoho, 2006]と呼ばれる技術に注目する。圧縮センシングは、少数の観測データから原信号を再構成することを目的としており、推定すべき信号がスパースであるという仮定に基づいている。[Kabashima et al., 2009; Ganguli and Sompolinsky, 2010]は、統計力学でモデルの分配関数を解析的に計算するために使われるレプリカ法[Nishimori, 2001]を用い、圧縮センシングによって原信号の適切な再構成が可能となる条件を理論的に解析した。その結果、データが少ない場合でも、適切な原信号の再構成に成功する領域が広く存在することが示された。

第３章では、上記の結果に着想を得て、ペアワイズマルコフ確率場の劣決定系における相互作用パラメータ推定の問題を、圧縮センシングで用いられる L1 ノルム最小化の形式で定式化する。さらに、レプリカ法を用いて推定の典型的な性能を理論的に解析する。また、この解析の妥当性を確かめるため、交互方向乗数法[Boyd et al., 2011]を用いて実際に未知の相互作用パラメータを推定する。レプリカ法による解析結果と数値実験による結果を比べ、解析結果をよく再現する結果が得られていることを確認する。この解析結果は、用意できるデータの数に対してどの程度までのパラメータ数であれば適切に推定を行うことができるか、または推定する必要のあるパラメータの数に対してどの程度のデータ数が必要なのかを示す。

第４章　平均場近似による推論

学習によってマルコフ確率場のパラメータが決定されれば、そのマルコフ確率場を用いて、未知のデータが与えられた場合のモデルの出力を推定することにより、予測や意思決定などを実現することができる。学習後のこの一連の処理は推論と呼ばれ、統計的機械学習モデルの実応用に

直結する重要な課題であると認識されている。マルコフ確率場は統計力学の基礎的な模型であるIsing 模型を一般化したモデルであることが知られている[Kindermann and Snell, 1980]。1990年代後半になると、Ising 模型の挙動を解析するために統計力学で発展してきた平均場近似と呼ばれる近似法が、統計的機械学習におけるマルコフ確率場の推論に転用され始めた[Tanaka, 1998; Kappen and Rodríguez, 1998]。これ以降、平均場近似は統計的機械学習モデルの近似推論に用いられる主要な方法として位置付けられ、多くの場面で利用されている[Opper and Saad, 2001; Hinton and Salakhutdinov, 2006]。

第４章では、Ising 変数（+1, -1 のいずれかの値をとる変数）を持つペアワイズマルコフ確率場に対する平均場近似の詳細な導出を示す。平均場近似の導出にはいくつか選択肢があるが、本論文ではそのうちの一つである、Plefka による Gibbs 自由エネルギーの摂動展開法[Plefka, 1982; Georges and Yedidia, 1991]を用いる。Gibbs 自由エネルギーは、マルコフ確率場の分配関数で表される量である Helmholz 自由エネルギーの双対変換として定義される。Gibbs 自由エネルギーを展開し、ナイーブ平均場近似および Thouless--Anderson--Palmer 近似を導出する。

第５章　多体相互作用を持つマルコフ確率場に対する発展的な平均場近似

第４章の説明で述べたように、平均場近似に基づく近似法はマルコフ確率場の推論に用いられる主要な計算法として知られている。最も基本的な平均場近似であるナイーブ平均場近似のほかに、Thouless--Anderson--Palmer 近似、Bethe 近似、適応 Thouless--Anderson--Palmer 近似などの発展的な平均場近似が用いられることもある。さらに、変数の高次の期待値を近似するために、このような平均場近似法とともに線形応答近似[Kappen and Rodriguez, 1998]と呼ばれる方法が用いられる。統計力学では、線形応答近似を用いて感受率（変数の二次のモーメント）を計算するメッセージ伝搬アルゴリズムである感受率伝搬法[Mézard and Mora, 2004]が用いられている。上記のような手法のペアワイズマルコフ確率場に対する適用については盛んに調べられているが、高次マルコフ確率場に対する適用についての研究は進んでいない。多体相互作用をモデルに含めることでモデルのパラメータとして陽に複数の変数の間の相互作用を表現することができるという利点はあるものの、計算処理が単純に複雑さを増すことから、高次マルコフ確率場に対する汎用的な推論アルゴリズムの開発は困難であるとして避けられている。

第５章では、高次マルコフ確率場に対する発展的な平均場近似の定式化を行う。まずはじめに、Ising 変数を持つペアワイズマルコフ確率場に対して提案された、改良された感受率伝搬法とナイーブ平均場近似を組み合わせた手法によって導出される近似方程式が、適応 Thouless--Anderson--Palmer 方程式と一致するという結果 [Yasuda and Tanaka, 2013] を拡張する。改良された感受率伝搬法とナイーブ平均場近似を組み合わせた手法を離散値または連続値の変数を持つペアワイズマルコフ確率場に対して適用した場合に導出される近似方程式が適応 Thouless--Anderson--Palmer 方程式と一致することを示す。ここで示した改良された感受率伝搬法とナイーブ平均場近似を組み合わせた手法と適応 Thouless--Anderson--Palmer 近似の等価性に基づいて、ペアワイズマルコフ確率場に対するナイーブ平均場近似と改良された感受率伝搬法を組み合わせた手法と同様の計算処理を行うことで、高次マルコフ確率場に対する発展的な平均場近似を

定式化する。定式化した提案手法はナイーブ平均場近似の拡張であるとみなすことができるため、ナイーブ平均場近似と提案手法の近似精度の数値的な比較を行い、提案手法の近似精度における優位性を示す。

第６章　結論

本論文では、確率的グラフィカルモデルの一つであるマルコフ確率場の数理モデル化に関する二つの研究結果を報告した。第３章ではペアワイズマルコフ確率場の学習、第４章および５章ではマルコフ確率場の推論にそれぞれ焦点を当て、統計力学の手法を用いた新しいアルゴリズムの提案および理論解析、数値実験を行った。第３章の研究からは、マルコフ確率場のパラメータの適切な推定に必要なデータ数の指標が得られた。例えばデータを得るためのコストが高い実験が必要な場合は、この指標に基づき、実験の回数を最小限に抑えながら適切なモデル化を行うことができる。第５章では、高次マルコフ確率場に対する高性能な平均場推論アルゴリズムを開発し、通常用いられる単純な平均場近似に比べて数値的に優れた性能を持つことを示した。